

#8

PCT09

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/857,583

DATE: 06/27/2001

TIME: 15:08:50

Input Set : A:\ES.txt

Output Set: N:\CRF3\06272001\I857583.raw

3 <110> APPLICANT: Browse, John et al.  
 5 <120> TITLE OF INVENTION: DESATURASES AND METHODS OF USING THEM FOR SYNTHESIS OF  
 POLYUNSATURATED

6 FATTY ACIDS

8 &lt;130&gt; FILE REFERENCE: 4630-58963

C--&gt; 10 &lt;140&gt; CURRENT APPLICATION NUMBER: US/09/857,583

C--&gt; 10 &lt;141&gt; CURRENT FILING DATE: 2001-06-05

10 &lt;150&gt; PRIOR APPLICATION NUMBER: US 60/111,301

11 &lt;151&gt; PRIOR FILING DATE: 1998-12-07

13 &lt;150&gt; PRIOR APPLICATION NUMBER: PCT/US99/28655

14 &lt;151&gt; PRIOR FILING DATE: 1999-12-06

16 &lt;160&gt; NUMBER OF SEQ ID NOS: 13

18 &lt;170&gt; SOFTWARE: PatentIn version 3.1

20 &lt;210&gt; SEQ ID NO: 1

21 &lt;211&gt; LENGTH: 1461

22 &lt;212&gt; TYPE: DNA

23 &lt;213&gt; ORGANISM: Caenorhabditis elegans

25 &lt;400&gt; SEQUENCE: 1

26	gaattttcaa	tcctccttgg	gtcccaccgc	tgtgatatca	aaatggtatt	acgagagcaa	60
28	gagcatgagc	cattcttcat	taaaattgat	ggaaaatggt	gtcaaattga	cgatgctgtc	120
30	ctgagatcac	atccaggtgg	tagtgcaatt	actacctata	aaaatatgga	tgccactacc	180
32	gtattccaca	cattccatag	tggttctaaa	gaagcgtatc	aatggctgac	agaattgaaa	240
34	aaagagtgcc	ctacacaaga	accagagatc	ccagatatta	aggatgacct	aatcaaagga	300
36	attgatgatg	tgaacatggg	aactttcaat	atttctgaga	aacgatctgc	ccaaataaat	360
38	aaaagtttca	ctgatctacg	tatgcgagtt	cgtgcagaag	gacttatgga	tggatctcct	420
40	ttgtttctaca	ttagaaaaat	tcttgaaaca	atcttcacaa	ttctttttgc	attctacctt	480
42	caataccaca	catattatct	tccatcagct	attctaattg	gagttgcgtg	gcaacaattg	540
44	ggatggttaa	tccatgaatt	cgcacatcat	cagttgttca	aaaacagata	ctacaatgat	600
46	ttggccagct	atttcgttgg	aaacttttta	caaggattct	catctggtgg	ttggaaagag	660
48	cagcacaatg	tgcatacgc	agccacaaat	gttgttgga	gagacggaga	tcttgattta	720
50	gtcccattct	atgctacagt	ggcagaacat	ctcaacaatt	attctcagga	ttcatgggtt	780
52	atgactctat	tcagatggca	acatgttcat	tggacattca	tgttaccatt	cctccgtctc	840
54	tcgtggcttc	ttcagtcaat	catttttgtt	agtcagatgc	caactcatta	ttatgactat	900
56	tacagaaata	ctgcgattta	tgaacagggt	ggtctctctt	tgcactgggc	ttggtcattg	960
58	ggtcaattgt	atttcttacc	cgattggtca	actagaataa	tgttcttctt	tgtttctcat	1020
60	cttggttgag	gtttcttgc	ctctcatgta	gttactttca	atcattatc	agtggagaag	1080
62	tttgcatgga	gctogaacat	catgtcaaat	tacgcttgtc	ttcaaatcat	gaccacaaga	1140
64	aatatgagac	ctggaagatt	cattgactgg	ctttggggag	gtcttaacta	tcagattgag	1200
66	caccatcttt	tcccacagat	gccacgacac	aacttgaaca	ctggttatgcc	acttggttaag	1260
68	gagtttgcag	cagcaaatgg	tttaccatac	atggtcgacg	attatttcac	aggattctgg	1320
70	cttgaaattg	agcaattccg	aaatattgca	aatgttgctg	ctaaattgac	taaaaagatt	1380
72	gcctagatta	cgattaatta	atcaatttat	tttcatgttc	tattcgtgtg	ttttaatatt	1440
74	ttccaaattt	ttacctattc	c				1461

77 &lt;210&gt; SEQ ID NO: 2

78 &lt;211&gt; LENGTH: 447

79 &lt;212&gt; TYPE: PRT

80 &lt;213&gt; ORGANISM: Caenorhabditis elegans

82 &lt;400&gt; SEQUENCE: 2

ENTERED

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/857,583

DATE: 06/27/2001

TIME: 15:08:50

Input Set : A:\ES.txt

Output Set: N:\CRF3\06272001\I857583.raw

```

84 Met Val Leu Arg Glu Gln Glu His Glu Pro Phe Phe Ile Lys Ile Asp
85 1 5 10 15
88 Gly Lys Trp Cys Gln Ile Asp Asp Ala Val Leu Arg Ser His Pro Gly
89 20 25 30
92 Gly Ser Ala Ile Thr Thr Tyr Lys Asn Met Asp Ala Thr Thr Val Phe
93 35 40 45
96 His Thr Phe His Thr Gly Ser Lys Glu Ala Tyr Gln Trp Leu Thr Glu
97 50 55 60
100 Leu Lys Lys Glu Cys Pro Thr Gln Glu Pro Glu Ile Pro Asp Ile Lys
101 65 70 75 80
104 Asp Asp Pro Ile Lys Gly Ile Asp Asp Val Asn Met Gly Thr Phe Asn
105 85 90 95
108 Ile Ser Glu Lys Arg Ser Ala Gln Ile Asn Lys Ser Phe Thr Asp Leu
109 100 105 110
112 Arg Met Arg Val Arg Ala Glu Gly Leu Met Asp Gly Ser Pro Leu Phe
113 115 120 125
116 Tyr Ile Arg Lys Ile Leu Glu Thr Ile Phe Thr Ile Leu Phe Ala Phe
117 130 135 140
120 Tyr Leu Gln Tyr His Thr Tyr Tyr Leu Pro Ser Ala Ile Leu Met Gly
121 145 150 155 160
124 Val Ala Trp Gln Gln Leu Gly Trp Leu Ile His Glu Phe Ala His His
125 165 170 175
128 Gln Leu Phe Lys Asn Arg Tyr Tyr Asn Asp Leu Ala Ser Tyr Phe Val
129 180 185 190
132 Gly Asn Phe Leu Gln Gly Phe Ser Ser Gly Gly Trp Lys Glu Gln His
133 195 200 205
136 Asn Val His His Ala Ala Thr Asn Val Val Gly Arg Asp Gly Asp Leu
137 210 215 220
140 Asp Leu Val Pro Phe Tyr Ala Thr Val Ala Glu His Leu Asn Asn Tyr
141 225 230 235 240
144 Ser Gln Asp Ser Trp Val Met Thr Leu Phe Arg Trp Gln His Val His
145 245 250 255
148 Trp Thr Phe Met Leu Pro Phe Leu Arg Leu Ser Trp Leu Leu Gln Ser
149 260 265 270
152 Ile Ile Phe Val Ser Gln Met Pro Thr His Tyr Tyr Asp Tyr Tyr Arg
153 275 280 285
156 Asn Thr Ala Ile Tyr Glu Gln Val Gly Leu Ser Leu His Trp Ala Trp
157 290 295 300
160 Ser Leu Gly Gln Leu Tyr Phe Leu Pro Asp Trp Ser Thr Arg Ile Met
161 305 310 315 320
164 Phe Phe Leu Val Ser His Leu Val Gly Gly Phe Leu Leu Ser His Val
165 325 330 335
168 Val Thr Phe Asn His Tyr Ser Val Glu Lys Phe Ala Leu Ser Ser Asn
169 340 345 350
172 Ile Met Ser Asn Tyr Ala Cys Leu Gln Ile Met Thr Thr Arg Asn Met
173 355 360 365
176 Arg Pro Gly Arg Phe Ile Asp Trp Leu Trp Gly Gly Leu Asn Tyr Gln
177 370 375 380
180 Ile Glu His His Leu Phe Pro Thr Met Pro Arg His Asn Leu Asn Thr

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/857,583

DATE: 06/27/2001

TIME: 15:08:50

Input Set : A:\ES.txt

Output Set: N:\CRF3\06272001\I857583.raw

```

181 385          390          395          400
184 Val Met Pro Leu Val Lys Glu Phe Ala Ala Ala Asn Gly Leu Pro Tyr
185          405          410          415
188 Met Val Asp Asp Tyr Phe Thr Gly Phe Trp Leu Glu Ile Glu Gln Phe
189          420          425          430
192 Arg Asn Ile Ala Asn Val Ala Ala Lys Leu Thr Lys Lys Ile Ala
193          435          440          445
196 <210> SEQ ID NO: 3
197 <211> LENGTH: 1275
198 <212> TYPE: DNA
199 <213> ORGANISM: Euglena gracilis
201 <400> SEQUENCE: 3
202 attttttttc gaaatgaagt caaagcgcca agcgctatcc cccttacaat tgatggaaca      60
204 aacatatgat gtggtcaatt tccaccctgg tgggtcggaa attatagaga attaccaagg      120
206 aagggatgcc actgatgcct tcattggttat gcactttcaa gaagccttcg acaagctcaa      180
208 gcgcatgccc aaaatcaatc ccagttttga gttgccaccc caggctgcag tgaatgaagc      240
210 tcaagaggat ttccggaagc tccgagaaga gttgatcgca actggcatgt ttgatgcctc      300
212 cccctctctg tactcataca aaatcagcac cacactgggc cttggagtgc tgggttattt      360
214 cctgatgggt cagtatcaga tgtatttcat tggggcagtg ttgcttgga tgcactatca      420
216 acagatgggc tggctttctc atgacatttg ccaccaccag actttcaaga accggaactg      480
218 gaacaacctc gtgggactgg tatttgga tggcttgcaa ggtttttccg tgacatgttg      540
220 gaaggacaga cacaatgcac atcattcggc aaccaatgtt caagggcacg accctgatat      600
222 tgacaacctc ccccccttag cctggtctga ggatgacgtc acacgggcgt caccgatttc      660
224 ccgcaagctc attcagttcc agcagtacta tttcttggtc atctgtatct tgttgcggtt      720
226 catttggtgt ttccagtgcg tgttgaccgt gcgcagtttg aaggacagag ataaccaatt      780
228 ctatcgtctc cagtataaga aggaggccat tggcctcgcc ctgcaactgga ccttgaaggc      840
230 cctgttccac ttattcttta tgcccagcat cctcacatcg ctggttggtgt ttttcgttct      900
232 ggagctgggt ggcggttcct gcattgcgat cgtggtgttc atgaaccact acccactgga      960
234 gaagatcggg gaccagctct gggatggcca tggattctcg gttggccaga tccatgagac      1020
236 catgaacatt cggcgaggga ttatcacaga ttggttttcc ggaggcttga attaccagat      1080
238 tgagcaccat ttgtggccga cctcctctcg ccacaacctg acagcggtta gctaccaggt      1140
240 ggaacagctg tgccagaagc acaacctgcc gtatcggaac ccgctgcccc atgaagggtt      1200
242 ggtcatcctg ctgcgctatc tggcgggtgt cgcccgatg gcggagaagc aaccgcggg      1260
244 gaaggtctta taagg      1275
247 <210> SEQ ID NO: 4
248 <211> LENGTH: 422
249 <212> TYPE: PRT
250 <213> ORGANISM: Euglena gracilis
252 <400> SEQUENCE: 4
254 Met Lys Ser Lys Arg Gln Ala Leu Ser Pro Leu Gln Leu Met Glu Gln
255 1          5          10          15
258 Thr Tyr Asp Val Ser Ala Trp Val Asn Phe His Pro Gly Gly Ala Glu
259          20          25          30
262 Ile Ile Glu Asn Tyr Gln Gly Arg Asp Ala Thr Asp Ala Phe Met Val
263          35          40          45
266 Met His Phe Gln Glu Ala Phe Asp Lys Leu Lys Arg Met Pro Lys Ile
267          50          55          60
270 Asn Pro Ser Phe Glu Leu Pro Pro Gln Ala Ala Val Asn Glu Ala Gln
271 65          70          75          80

```

## RAW SEQUENCE LISTING

DATE: 06/27/2001

PATENT APPLICATION: US/09/857,583

TIME: 15:08:50

Input Set : A:\ES.txt

Output Set: N:\CRF3\06272001\I857583.raw

```

274 Glu Asp Phe Arg Lys Leu Arg Glu Glu Leu Ile Ala Thr Gly Met Phe
275      85      90      95
278 Asp Ala Ser Pro Leu Trp Tyr Ser Tyr Lys Ile Ser Thr Thr Leu Gly
279      100     105     110
282 Leu Gly Val Leu Gly Tyr Phe Leu Met Val Gln Tyr Gln Met Tyr Phe
283      115     120     125
286 Ile Gly Ala Val Leu Leu Gly Met His Tyr Gln Gln Met Gly Trp Leu
287      130     135     140
290 Ser His Asp Ile Cys His His Gln Thr Phe Lys Asn Arg Asn Trp Asn
291 145      150     155     160
294 Asn Leu Val Gly Leu Val Phe Gly Asn Gly Leu Gln Gly Phe Ser Val
295      165     170     175
298 Thr Cys Trp Lys Asp Arg His Asn Ala His His Ser Ala Thr Asn Val
299      180     185     190
302 Gln Gly His Asp Pro Asp Ile Asp Asn Leu Pro Pro Leu Ala Trp Ser
303      195     200     205
306 Glu Asp Asp Val Thr Arg Ala Ser Pro Ile Ser Arg Lys Leu Ile Gln
307      210     215     220
310 Phe Gln Gln Tyr Tyr Phe Leu Val Ile Cys Ile Leu Leu Arg Phe Ile
311 225      230     235     240
314 Trp Cys Phe Gln Cys Val Leu Thr Val Arg Ser Leu Lys Asp Arg Asp
315      245     250     255
318 Asn Gln Phe Tyr Arg Ser Gln Tyr Lys Lys Glu Ala Ile Gly Leu Ala
319      260     265     270
322 Leu His Trp Thr Leu Lys Ala Leu Phe His Leu Phe Phe Met Pro Ser
323      275     280     285
326 Ile Leu Thr Ser Leu Leu Val Phe Phe Val Ser Glu Leu Val Gly Gly
327      290     295     300
330 Phe Gly Ile Ala Ile Val Val Phe Met Asn His Tyr Pro Leu Glu Lys
331 305      310     315     320
334 Ile Gly Asp Pro Val Trp Asp Gly His Gly Phe Ser Val Gly Gln Ile
335      325     330     335
338 His Glu Thr Met Asn Ile Arg Arg Gly Ile Ile Thr Asp Trp Phe Phe
339      340     345     350
342 Gly Gly Leu Asn Tyr Gln Ile Glu His His Leu Trp Pro Thr Leu Pro
343      355     360     365
346 Arg His Asn Leu Thr Ala Val Ser Tyr Gln Val Glu Gln Leu Cys Gln
347      370     375     380
350 Lys His Asn Leu Pro Tyr Arg Asn Pro Leu Pro His Glu Gly Leu Val
351 385      390     395     400
354 Ile Leu Leu Arg Tyr Leu Ala Val Phe Ala Arg Met Ala Glu Lys Gln
355      405     410     415
358 Pro Ala Gly Lys Ala Leu
359      420
362 <210> SEQ ID NO: 5
363 <211> LENGTH: 27
364 <212> TYPE: DNA
365 <213> ORGANISM: artificial sequence
367 <220> FEATURE:

```

RAW SEQUENCE LISTING  
 PATENT APPLICATION: US/09/857,583

DATE: 06/27/2001  
 TIME: 15:08:50

Input Set : A:\ES.txt  
 Output Set: N:\CRF3\06272001\I857583.raw

```

368 <223> OTHER INFORMATION: PCR Primer
370 <220> FEATURE:
371 <221> NAME/KEY: misc_feature
372 <222> LOCATION: (12)..(12)
373 <223> OTHER INFORMATION: n = a, t, g, or c
376 <400> SEQUENCE: 5
W--> 377 ggctggctga cncaygartt ytgycay 27
380 <210> SEQ ID NO: 6
381 <211> LENGTH: 30
382 <212> TYPE: DNA
383 <213> ORGANISM: artificial sequence
385 <220> FEATURE:
386 <223> OTHER INFORMATION: PCR Primer
388 <220> FEATURE:
389 <221> NAME/KEY: misc_feature
390 <222> LOCATION: (13)..(13)
391 <223> OTHER INFORMATION: n = a, t, g, or c
394 <400> SEQUENCE: 6
W--> 395 catcgttggga aanarrtgrt gytcdatytg 30
398 <210> SEQ ID NO: 7
399 <211> LENGTH: 41
400 <212> TYPE: DNA
401 <213> ORGANISM: artificial sequence
403 <220> FEATURE:
404 <223> OTHER INFORMATION: PCR Primer
406 <400> SEQUENCE: 7
407 cccgggaagc ttctcgagga attttcaatc ctccttgggt c 41
410 <210> SEQ ID NO: 8
411 <211> LENGTH: 34
412 <212> TYPE: DNA
413 <213> ORGANISM: artificial sequence
415 <220> FEATURE:
416 <223> OTHER INFORMATION: PCR Primer
418 <400> SEQUENCE: 8
419 cccgggtgga tccggaacat atcacacgaa acag 34
422 <210> SEQ ID NO: 9
423 <211> LENGTH: 6
424 <212> TYPE: RNA
425 <213> ORGANISM: artificial sequence
427 <220> FEATURE:
428 <223> OTHER INFORMATION: Consensus sequence
430 <220> FEATURE:
431 <221> NAME/KEY: polyA_signal
432 <222> LOCATION: (1)..(6)
433 <223> OTHER INFORMATION:
436 <400> SEQUENCE: 9
437 aauaaa 6
440 <210> SEQ ID NO: 10
441 <211> LENGTH: 20

```

Use of n and/or Xaa has been detected in the Sequence Listing.  
 Review the Sequence Listing to insure a corresponding  
 explanation is presented in the <220> to <223> fields of  
 each sequence using n or Xaa.

## VERIFICATION SUMMARY

PATENT APPLICATION: US/09/857,583

DATE: 06/27/2001

TIME: 15:08:51

Input Set : A:\ES.txt

Output Set: N:\CRF3\06272001\I857583.raw

L:10 M:270 C: Current Application Number differs, Replaced Current Application No  
L:10 M:271 C: Current Filing Date differs, Replaced Current Filing Date  
L:377 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5  
L:395 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6  
L:487 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:12  
L:508 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13